Cluster-randomized controlled trials of individual and combined water, sanitation, hygiene, and nutritional interventions in rural Bangladesh and Kenya: The WASH Benefits Study design and rationale

Appendix 5. Pre-specified secondary analyses: treatment interactions, cross-cluster externalities, subgroup analyses.

Tests of treatment interaction

The study is powered for the tests described in the main text. We chose to design the study around main effects and not these interaction tests because we expect the interactions, if present, to be small and thus difficult to detect in feasible designs. However, the design will enable us to test for large interactions between treatments (related to H2 and H3 in the main text). The rationale for including the interaction tests in our analysis plan is that if the interactions are large, they will be both detectable and scientifically important. Nonetheless, we recognize that the study will not have power to detect these interactions unless they are at least 2 times larger than the main effects.[1] This is because the interaction tests will rely on variance terms from more than 2 arms (in contrast to the parameters described in the main text). The interactions we describe below are on the additive scale.

The first interaction test is whether combined water quality, handwashing, and sanitation interventions improve our primary outcomes more (or less) than the additive effect of the components delivered separately. The The null hypothesis is:

$$Ho:\ E_Z(\ [Y \mid T = WSH, Z\ ]\ ) = E_Z\ (\ \mathrm{E}[Y \mid T = W, Z\ ] + \mathrm{E}[Y \mid T = S, Z\ ] + \mathrm{E}[Y \mid T = H, Z\ ]\ )$$

There are theoretical [2–4] and observational [5,6] studies to support this hypothesis, but the only randomized trial to date found no positive interaction between water treatment and handwashing [7] (and, if anything, antagonism, where the effect of the combined treatment is less than the additive effect of water treatment + handwashing).

The second interaction test is whether combined WASH and Nutrition interventions improve our primary outcomes more (or less) than the additive effect of the components delivered separately. The null hypothesis is:

$$Ho:\ E_Z(\ \mathrm{E}[Y \mid T = NWSH, Z\ ]\ ) = E_Z\ (\ \mathrm{E}[Y \mid T = WSH, Z\ ] + \mathrm{E}[Y \mid T = N, Z\ ]\ )$$

While there is biologic plausibility for this interaction, there is scant empirical evidence to support or refute the hypothesis.[8]

Testing for- and estimating cross-cluster spillovers/externalities

A fundamental assumption for unbiased causal inference in a randomized trial is that the units of randomization are independent.[9] In this study, clusters are the unit of randomization. Cross-cluster spillover effects occur when the treatment assignment of one cluster influences outcomes in another cluster. The mechanism for spillover could be through disease transmission or through information diffusion. Although we expect a priori that cross-cluster spillovers are likely to be quite small, we plan to test this assumption. We plan to test for spillovers over geographic distance and through shared

school membership and market attendance. In the notation below we use distance as an example, but the same parameter and notation applies to spillovers through other channels, which we will test for separately.

Let N_d^T be the number of treated compounds with treatment *T* in some distance *d*, defined as straight-line, geographic distance from the cluster perimeter. We do not control N_d^T by design, but we expect that there will be random variation created by our design. Define a new parameter among the control clusters (*T*=*c*), which includes the effect of adjacent treated compounds (N_d^T) as a measure of the spillover effect, controlling for cluster-level covariates *X*:

$$\theta = E_{X,N^T}\left(E[Y|T=c,X,N_d^T] - E[Y|T=c,X,N_d^T=0]\right)$$

To estimate this parameter, we will need to model E[*Y* | *T, X,* N_d^T]. The linear model used by Miguel and Kremer [10] is a sensible choice, but we may consider less parametric prediction algorithms.[11] To test for cross-cluster spillovers, we will restrict the analysis to the control clusters to simplify the test. The first term is the empirical distribution of *Y* in the control group, including observed spillovers (N_d^T). The second term is estimated from the predicted values of *Y* from the algorithmic fit under conditions of no spillover effects (N_d^T =0). (Note: if there are no clusters without spillover effects, the model would need to extrapolate beyond the observed data.) Under the null hypothesis of no spillovers, the parameter equals zero. The null hypothesis is:

$$Ho: Y \perp N_d^T \mid T, X$$

We can test the null hypothesis with a clustered permutation test for each treatment, *T*. This involves permuting the cluster IDs in the control group, re-fitting the algorithm, and re-estimating *θ* for a large number of permutations. This will generate a null distribution of *θ*. We can then obtain a *P*-value for the test by comparing the observed *θ* to its null distribution.

If we cannot reject the null hypothesis, then we will proceed with the standard Intention-To-Treat (ITT) analysis (parameters described in the main text). If we reject the null hypothesis, then *θ* will provide an estimate of the magnitude of spillover effects for each treatment *T*. In the presence of spillovers the ITT estimates will be a lower bound of the estimate of the total effect of treatment under the assumption that spillover effects are positive.

Scope: We plan to test for spillovers in behavior change uptake indicators (Appendix 2) and our primary outcomes. We will repeat the test for each outcome and treatment. We do not expect spillover effects from the nutrition intervention treatment and will not test for them. We will test for spillovers through three main channels:

1. Geographic proximity, with bands (*d*) similar to Miguel and Kremer [10] defined after the baseline survey (not using outcomes) when we have a sense for relevant geographic distances between clusters in each country
2. School attendance
3. Market attendance

To help improve the estimation in all cases, we will attempt collect some measure of total population or compounds in each institution as a variable in *X* to control for differences in density.

Pre-specified subgroup analyses

We recognize that the study is powered to detect main effects on our primary outcomes, and so we will be unlikely to detect subgroup-specific effects unless they are larger than the overall ITT effect.[1] However, we feel that some of the subgroup-specific effects are highly relevant to interpreting the study and to informing intervention targeting in the future. This type of analysis extends the interaction tests between treatments described above by looking at treatment interactions with baseline covariates. For example, the most relevant effect of a water quality intervention is among households who have poor water quality at baseline; it is less likely that a water quality intervention would improve health among children who live in households with microbiologically clean drinking water at baseline.

For all of the subgroup-specific effects that that we plan estimate *a priori* in this study, we will first screen the variables to ensure that there is sufficient variation for the tests to make sense. We will estimate different ITT effects for the different subgroups by interacting subgroup variables with the treatment indicators of interest.[1,12] Within each category of baseline covariates, the country teams have selected characteristics that they will include in subgroup analyses.

<u>Household water treatment and quality, source water access and water quality</u>

Rationale: The effect of our drinking water quality intervention may be smaller among households with good baseline drinking water quality. The effect of our other WASH interventions may be greater or smaller, depending on baseline drinking water quality and water source availability. In Kenya, we expect that the majority of our study population will have received a Lifestraw family filter as part of a Vestergaard Frandsen (VF) distribution program throughout Western Province. If the filters are in regular use, we would expect smaller impacts from the chlorine dispenser intervention among those households.

Both countries

- Drinking water source (surface water vs. other)
- Household reports regularly treating their drinking water
- Free residual chlorine in stored drinking water

Kenya

- Detectable *E. coli* in source water (> 0 CFU / 100 ml)
- Detectable *E. coli* in drinking water (> 0 CFU / 100 ml)
- Field staff observe a VF water filter hanging in the household and household members report frequent use
- Observed VF water filter has visible moisture in it.
- Walking distance in minutes to primary drinking water source

<u>Handwashing practices</u>

Rationale: The effect of our handwashing intervention may be smaller among households with good baseline handwashing practices. The effect of our other WASH interventions may be greater or smaller, depending on baseline handwashing practices.

Both countries
- Mother has clean palms, finger pads, and finger nails

Kenya
- Mother was observed to use soap during a handwashing demonstration
- Mother lists (unprompted) as critical times for handwashing: before preparing food, eating, or feeding a child and after defecating or cleaning a child who has defecated.

Bangladesh
- Presence of a handwashing station with water and soap

Sanitation conditions

Rationale: The effect of our sanitation intervention may be smaller among households with high levels of baseline sanitation. The effect of our other WASH interventions may be greater or smaller, depending on baseline sanitation conditions. For example, an observational study using DHS data documented larger effects of improved source water only in the presence of improved sanitation conditions.[6]

Both countries
- Household latrine status (none, unimproved, JMP improved)

Kenya
- Stool visible on floor of the latrine
- Any person in household reported to not always use latrine
- Most recent feces of child under 36 months were disposed of in latrine
- Latrine is located in another compound
- Household already owns potty
- Cover observed over latrine drop hole

Food security

Rationale: The effect of our Nutrition intervention or combined Nutrition+WSH intervention may be greater among households with low food security at baseline.

Both countries
- Questions will be adapted from the Household Food Insecurity and Access Scale (HFIAS), with modifications for the local language, cultural context, and food availability patterns.

Child age

Rationale: All target children will be enrolled in the study while in utero, but their experience of the intervention will differ slightly depending on their relative age within the cohort, which will span approximately 6 months of age. Our outcome measurements will

take place at a fixed calendar time – not child age. It is possible that younger children will benefit more from being born into more mature intervention conditions. A competing hypothesis is that the younger children will benefit less from intervention because they will have had less post-natal exposure compared to older children.

Both countries
- Stratify the results by age in 3-month brackets at the endline survey: [18,21), [21, 24), [24, 27)

Child sex

Rationale: Biologic differences, cultural practices, or behavioral practices may modify the effect of the interventions with respect to boys or girls.

Both countries
- Stratify the results by sex

Number of older children living in the compound

Rationale: Children living in compounds with older children may be at higher risk for pathogen transmission into the compound. Older children have greater exposure through schools and social networks, and if they do not use latrines they may have greater pathogen shedding in the compound through open defecation.

Both countries
- Stratify the results by the number of older children (<15 years old) in the compound.

Cluster density and cluster size

Rationale: The positive or negative effects of proximate neighbors may modify the protective effects of the intervention. For negative spillover effects, like disease transmission, we would expect the interventions to be less efficacious in densely populated environments than in more sparsely populated environments. In Kenya, where cluster sizes vary, is possible that the intervention effects will be heterogeneous with cluster size because the number of treated households per intervention promoter may change the nature of the promotion.

Both countries
- Stratify the results into clusters of high compound density and low compound density.

Kenya only
- Stratify the results by the number of households per promoter in the cluster.

Maternal intelligence and education

Rationale: mothers who are better educated and/or perform better on literacy tests may be more capable of adapting to new information and technology. They may have greater ability to optimize their behavior to take advantage of the messages and materials that the study provides.

Both countries

- Mothers that score in the top 25th percentile of the study population on at least one of the maternal intelligence tests that we administer at the 1-year follow-up survey
- Maternal schooling attainment

Kenya

- Maternal self-reported literacy

Appendix References

1 Brookes ST, Whitely E, Egger M, *et al.* Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol* 2004;**57**:229–36.

2 Briscoe J. Intervention studies and the definition of dominant transmission routes. *Am J Epidemiol* 1984;**120**:449–55.

3 VanDerslice J, Briscoe J. Environmental interventions in developing countries: interactions and their implications. *Am J Epidemiol* 1995;**141**:135–44.

4 Eisenberg JNS, Scott JC, Porco T. Integrating disease control strategies: balancing water sanitation and hygiene interventions to reduce diarrheal disease burden. *Am J Public Health* 2007;**97**:846–52.

5 Esrey SA, Habicht JP, Casella G. The complementary effect of latrines and increased water usage on the growth of infants in rural Lesotho. *Am J Epidemiol* 1992;**135**:659–66.

6 Esrey SA. Water, waste, and well-being: a multicountry study. *Am J Epidemiol* 1996;**143**:608–23.

7 Luby SP, Agboatwalla M, Painter J, *et al.* Combining drinking water treatment and hand washing for diarrhoea prevention, a cluster randomised controlled trial. *Trop Med Int Health* 2006;**11**:479–89.

8 Dewey KG, Mayers DR. Early child growth: how do nutrition and infection interact? *Matern Child Nutr* 2011;**7 Suppl 3**:129–42.

9 Holland PW. Statistics and Causal Inference. *J Am Stat Assoc* 1986;**81**:945–60.

10 Miguel E, Kremer M. Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica* 2004;**72**:159–217.

11 Van der Laan MJ, Polley EC, Hubbard AE. Super learner. *Stat Appl Genet Mol Biol* 2007;**6**:Article 25.

12 Pocock SJ, Assmann SE, Enos LE, *et al.* Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Stat Med* 2002;**21**:2917–30.